FUNDAÇÃO
GETULIO VARGAS

EPGE

Escola de Pós-Graduação
em Economia

Ensaios Econômicos

Escola de
Pós-Graduação
em Economia
da Fundação
Getulio Vargas

N° 453 ISSN 0104-8910

On Certain Geometric Aspects of Portfolio Optimisation with Higher Moments

Renato Galvão Flôres Junior, Gustavo Monteiro de Athayde

Agosto de 2002

URL: http://hdl.handle.net/10438/435

Galvão Flôres Junior, Renato
On Certain Geometric Aspects of Portfolio Optimisation with Higher Moments/ Renato Galvão Flôres Junior, Gustavo Monteiro de Athayde - Rio de Janeiro : FGV,EPGE, 2010
(Ensaios Econômicos; 453)

Inclui bibliografia.

CDD-330

On Certain Geometric Aspects of Portfolio Optimisation with Higher Moments*

Gustavo M. de Athayde and *Renato G. Flôres Jr.*
Banco Itaú S. A., São Paulo EPGE/Fundação Getulio Vargas, Rio

ABSTRACT

We discuss geometric properties related to the minimisation of a portfolio kurtosis given its first two odd moments, considering a risk-less asset and allowing for short sales. The findings are generalised for the minimisation of any given even portfolio moment with fixed excess return and skewness, and then for the case in which only excess return is constrained. An example with two risky assets provides a better insight on the problems related to the solutions. The importance of the geometric properties and their use in the higher moments portfolio choice context is highlighted.

JEL classification: C49; C61; C63.

(This version: August 5, 2002)

* This is a paper specially prepared for the *Multi-Moment Capital Asset Pricing Models and Related Topics Workshop* held in Paris on April 29, 2002. We thank all the participants for comments and lively related discussions and, particularly, Emmanuel Jurczenko and Bertrand Maillet, the workshop organisers on behalf of the *Finance-sur-Seine Association.* All ideas, affirmatives and mistakes are our own.

1. Introduction.

Portfolio optimisation taking into account more than the first two moments has been receiving renewed interest in the past years. Be it on the theoretical side – including its links with the CAPM extensions -, or on what relates to econometric tests or updates based on higher conditional moments, works like Adcock (2002), Adcock and Shutes (1999), Athayde and Flôres (1997, 2001, 2002), Jurczenko and Maillet (2001, 2002), Pedersen and Satchell (1998), or Athayde and Flôres (2000), Barone-Adesi (1985), Harvey and Siddique (1999, 2000), Hwang and Satchell (1999) and Pedersen and Satchell (2000), far from exhausting the full list of contributions, pay good witness to the growing awareness of the importance of higher moments in both lines of research.

Since Athayde and Flôres (1997), we have developed a systematic way to treat the key optimisation problems posed to anyone dealing with higher moments in portfolio theory. The approach uses a new notation to represent any moments tensor related to a multivariate random vector of asset returns, and can be used either in a utility maximising context or if optimal portfolios are defined by preference relations. The new notation seemed necessary in order to treat the problem in an absolutely general setting, which means both in the maximum order p of portfolio moments of interest and in the possible patterns of the skewness or higher order tensors. The latter is crucial as many works generalising the subject consider only the *marginal* higher moments of the returns vector, plainly disregarding any co-moment of the same order. Though the full set of co-moments can quickly become prohibitive – what, beyond other questions, may pose serious econometric estimation problems for the applications -, and simplifying hypotheses on its pattern will usually be imposed in practice, it is important to have a way to study the general solution to the problem, *irrespective of further assumptions that might be imposed.*

The utility function approach, given its more rigid theoretical constraints, and the debates involving any non-normality-implying (utility) function proposed, seems more suitable for theoretical developments related, for instance, to the CAPM. Preference ordering of portfolios, made rigorous by Scott and Horvath (1980), can lead to more interesting results in the strict portfolio optimisation context.

In this paper, we discuss an interesting geometric structure that arises when optimising an *even* moment subject to *odd* moments constraints. As usual, agents "like" odd moments and "dislike" even ones.

The structure studied – not the only relevant one in the higher moments context – bears important consequences and sheds light on the geometry of efficient portfolios sets in moments space. We believe that its implications have not been fully exploited yet. Moreover, final testing of the gains brought out by using higher moments relies in extensive practical applications of the new results. These, in turn, require proper software tools for solving the non-linear systems and optimisation problems involved. Better knowledge of the surfaces (or manifolds) related to them may greatly improve the software design.

This paper is organised as follows. The next section discusses the optimisation of variance, and then kurtosis, given the first and third desired portfolio moments; while section three discusses how these results could still be generalised. Section four draws, through an example, a few more properties and analyses the sensitivity of certain solutions. The final section concludes by explaining how the results can be useful in a duality context and sets a few lines of further research. An Appendix provides a brief explanation of the notation used.

2. Minimal variances and kurtoses subject to the first two odd moments.

Even moments, being always non-negative, are duly associated with spread, and both variance and kurtosis are used as simple numerical summaries of the dispersion of a set of observations. For fixed portfolio return and skewness, the latter should perhaps be more used in practice as an alternative objective function, given the frequency with which the fat-tailed effect in stock returns has been detected. If we minimise the fourth moment, we shall be directly attacking the heavy extremes of the density, the ultimate culprits of the high volatility and uncertainty of returns. Most measures of risk focused on the worst scenarios, like the VaR, would probably be more sensitive to variations in the fourth moment rather than in variance. This sort of behaviour will be further examined in the example in section 4. The material in this section draws on parts of Athayde and Flôres (2001) – where a complete solution to the three moments portfolio problem is found – and Athayde and Flôres (2002), for the developments related to the kurtosis; proofs omitted here can be found in these papers.

2.1. Homothetic properties of the minimum variance set.

Minimising the variance, for a given mean return and skewness, amounts to find the solution to the problem:

$$Min_{\alpha} L = \alpha' M_2 \alpha + \lambda_1 [E(r_p) - (\alpha' M_1 + (1 - \alpha' [1]) r_f)] + \lambda_2 (\sigma_{p^3} - \alpha' M_3 (\alpha \otimes \alpha)) \quad , \quad (1)$$

where M_1 , M_2 and M_3 are, resp., the matrices related to the first, second and third moments tensors[1], α is the vector of n portfolio weights – where short sales are allowed, r_f is the riskless rate of return, [1] stands for a nx1 vector of 1's, the lambdas are Lagrange multipliers and the two remaining symbols are the α-portfolio (given) mean return and skewness.

Calling $x = M_1 - [1]\, r_f$, the vector of mean excess returns, and

$R = E(r_p) - r_f$ the set (excess) portfolio return,

the solution to (1) is found by solving the n-equations non-linear system,

$$M_2 \alpha = \frac{A_4 R - A_2 \sigma_{p^3}}{A_0 A_4 - (A_2)^2} x + \frac{A_0 \sigma_{p^3} - A_2 R}{A_0 A_4 - (A_2)^2} M_3 (\alpha \otimes \alpha) \quad . \quad (2)$$

where the scalars:

$$A_0 = x' M_2^{-1} x \quad ,$$

$$A_2 = x' M_2^{-1} M_3 (\alpha \otimes \alpha) \quad , \quad (3)$$

$$A_4 = (\alpha \otimes \alpha)' M_3 M_2^{-1} M_3 (\alpha \otimes \alpha) \quad ,$$

have subscripts corresponding to their degree of homogeneity as (real) functions of the vector α. A_0 and A_4, in particular, are positive because the inverse of the covariance matrix is positive definite.

Pre-multiplying (2) by the very solutions α', gives the optimal variance(s):

$$\sigma_{p^2} = \frac{A_4 R^2 - 2 A_2 \sigma_{p^3} R + A_0 (\sigma_{p^3})^2}{A_0 A_4 - (A_2)^2} \quad , \quad (4)$$

an expression where both the numerator and denominator are positive.

The following proposition is fundamental:

Proposition 1: For a given *k*, let $\overline{\alpha}$ define the minimum variance portfolio when *R=1* and $y_3 = \sqrt[3]{\sigma_{p^3}} = k$, and $\overline{\sigma}_{p^2}$ be the corresponding minimum variance, THEN for all optimal portfolios related to return and skewness pairs (R, σ_{p^3}) such that $\sigma_{p^3} = k^3 R^3$, or $y_3 = kR$, the solution to (1) will be $\alpha = \overline{\alpha} R$, with corresponding minimum variance $\sigma_{p^2} = \overline{\sigma}_{p^2} R^2$.

The above result implies that along the direction defined in the returns x skewness plan by $y_3 = kR$, the optimal variance as a function of the excess return will be a parabola. Taking now the three dimensional (3D) space where the standard deviation $\sqrt[2]{\sigma_{p^2}} = y_2$ axis is added, in the half-plane formed by a specific direction *k* in R x y3 space[2] and the positive part of the standard deviation axis, the optimal portfolio surface will be reduced to the straight line $y_{p^2} = \overline{y}_{p^2} \dfrac{u}{\sqrt{k^2+1}}, \quad u \geq 0$[3]. As $\overline{y}_{p^2}$ differs with *k*, the angle that this line makes with the standard deviation axis varies also with *k*,.

The Proposition has then a far reaching consequence: the optimal surface in the positive standard deviation (sd) half of 3D space bears a homothetic property from whatever standpoint one assumes. Slicing the surface by a sequence of planes parallel to the two odd-moments axes will generate a sequence of curves starting at the origin and whose expansion ratio will be equal to that of the respective (constant) variance values. Of course, slicing it by planes parallel to the sd and (standardised) skewness axes will

---

[1] See the Appendix for a further explanation on the notation used.

[2] We shall, from now on, use the angular coefficient *k* to name the corresponding line/direction in the first quadrant of the R x y3 plane.

produce a sequence of homothetic curves whose expansion ratio will be that of the (excess) returns associated to each plan. Finally, inspection of formula (4) easily convinces that for the last combination, i.e. planes parallel to the sd and mean returns axes, the same will apply, as Proposition 1 is also true if the role of returns and skewness are reversed.

The proposition below is a direct consequence of this important fact:

Proposition 2: For a given level of y_2 (or R, or y_3), cut the optimal surface with a plane orthogonal to the sd (or returns, or standardised skewness) axis and project the intersection curve in the 'returns x skewness (or sd x skewness, or returns x sd) plane',

THEN

if they exist, the directions in the R x y_3 (or y_2 x y_3, or R x y_2) half plane related to the highest and lowest value, in each axis, of the curve are invariant with y_2 (or R, or y_3).

The qualification *if they exist* is important as, specially in the case of cuts parallel to the sd axis, at least part of the curve may go to infinity. For constant variance cuts, it may be shown that closed curves will be produced[4]. Indeed, for this case, the highest and lowest directions are particularly noteworthy, as demonstrated by

Proposition 3: The direction in the R x y_3 half plane that gives the highest R for all the minimum variance portfolios with the same standard deviation y_2 is unique and related to the celebrated (Markowitz's) Capital Market Line (CML). Moreover, in this direction, the skewness constraint to programme (1) is not binding. As regards

[3] The variable *u* stands for the coordinates along the axis defined by the "direction k".
[4] The proof is rather technical to be included in this text.

skewness, though there may be more than one "highest" (and "lowest") direction, the constraint property also applies.

This means that the unique solution to the minimum variance portfolio, for a given mean return:

$$\alpha = \frac{R}{A_0} M_2^{-1} x \qquad , \tag{5}$$

that defines the famous *Capital Market Line* in mean x variance space, relating the optimal variance to the given R,

$$\sigma_{R^2} = \frac{R^2}{A_0} \qquad , \tag{6}$$

also defines the (unique) direction that will pass through all the points, in each curve, yielding the maximum mean return. In other words, in the R x y_3 half plane, this direction is the geometric locus of all the tangency points between each (projected) curve and a straight line, parallel to the skewness axis, which cuts the mean return axis in the *maximum mean return portfolio value related to the set variance (that defines the cut).* This last statement is ensured by the well-known duality result in Markowitz world.

Skewnesses - and a k - can also be associated to these optimal portfolios, being evident that they are independent of the given y_2. It can be proved that the k – the angular coefficient of the line related to the extreme means – will be equal to:

$$k_R = \frac{y_{R^3}}{R} = \frac{\sqrt[3]{w' M_3 (w \otimes w)}}{A_0} \qquad , \qquad \text{where} \quad w = M_2^{-1} x \quad . \tag{7}$$

Hence, k_R is indeed an invariant and all maximum mean returns for given variances lie in the same direction in mean x skewness space.

Contrary to the previous, mean returns case, the optimal weights for the skewness extremes are implicitly defined by a non-linear system like (2). When $\sigma_{p^3} = 1$, we have a solution portfolio $\overline{\alpha}_S$ such that:

$$\overline{\alpha}_s = \frac{1}{\overline{A}_4} M_2^{-1} M_3 (\overline{\alpha}_S \otimes \overline{\alpha}_S) \quad . \tag{8}$$

The homothecy implies that $\alpha_s = \overline{\alpha}_s \sqrt[3]{\sigma_{s^3}} = \overline{\alpha}_s y_{s^3}$ is a solution to (2), ensuing an optimal variance $\sigma_{s^2} = \overline{\sigma}_{s^2} (y_{s^3})^2$. A corresponding (excess) return and a direction, both independent of the variance level, can be found as:

$$R_s = \overline{R}_s y_{p^3} \quad , \quad k_S = 1/\overline{R}_S \quad , \tag{9}$$

implying that all these optimal portfolios lie in the same direction.

Combining both results gives a rectangular envelope that circumscribes, in the first quadrant of the mean x skewness plane, the corresponding part of the constant variance curve.

2.2. The minimum kurtosis case.

The initial step now is minimising kurtosis for a given skewness and expected return:

$$Min_\alpha \; \alpha' M_4 (\alpha \otimes \alpha \otimes \alpha) + \lambda [E(r_p) - (\alpha' M_1 + (1 - \alpha' [1]) r_f)] + \gamma (\sigma_{p^3} - \alpha' M_3 (\alpha \otimes \alpha)) \; . \tag{10}$$

The first order conditions are:

$$\begin{aligned} &4 M_4 (\alpha \otimes \alpha \otimes \alpha) = \lambda x + 3 \gamma M_3 (\alpha \otimes \alpha) \\ &R = E(r_p) - r_f = \alpha' (M_1 - r_f [1]) = \alpha' x \\ &\sigma_{p^3} = \alpha' M_3 (\alpha \otimes \alpha) \end{aligned} \tag{11}$$

Defining

$$B_{(-2)} = x' \left[M_4(\alpha \otimes \alpha \otimes I) \right]^{-1} x \quad ,$$

$$B_0 = x' \left[M_4(\alpha \otimes \alpha \otimes I) \right]^{-1} M_3(\alpha \otimes \alpha) \quad \text{and}$$

$$B_2 = (\alpha \otimes \alpha)' M_3' \left[M_4(\alpha \otimes \alpha \otimes I) \right]^{-1} M_3(\alpha \otimes \alpha) \quad ,$$

with the subscripts chosen according to the degree of homogeneity of the term with respect to the vector α , one can find the values of λ and γ and arrive at the non-linear system that characterises the solution to (10):

$$M_4(\alpha \otimes \alpha \otimes \alpha) = \frac{B_2 R - B_0 \sigma_{p^3}}{B_{(-2)} B_2 - (B_0)^2} x + \frac{B_{(-2)} \sigma_{p^3} - B_0 R}{B_{(-2)} B_2 - (B_0)^2} M_3(\alpha \otimes \alpha) \quad . \qquad (12)$$

The optimal kurtosis will be given by:

$$\sigma_{p^4} = \frac{B_2 R^2 - 2 B_0 R \sigma_{p^3} + B_{(-2)} (\sigma_{p^3})^2}{B_{(-2)} B_2 - (B_0)^2} \quad . \qquad (13)$$

Noticing that $B_{(-2)}$ and B_2 are positive, because the matrix in their middle is the inverse of a positive definite matrix, it can be proved that both the numerator and the denominator of the expression above are positive.

It is important to remark the similarities between the pairs of formulas (2)-(12) and (4)-(13), as they are at the heart of the similar developments that follow. The first is a key proposition, close to Proposition 1:

Proposition 1*: For a given k, all the minimum kurtosis portfolios related to expected returns, skewnesses pairs (R, σ_{p^3}) such that $\sigma_{p^3} = k^3 R^3$, or $y_3 = kR$, are given by $\alpha = \overline{\alpha} R$, where $\overline{\alpha}$ defines $\overline{\sigma}_{p^4}$, the (minimum) kurtosis of the optimal portfolio when R=1 and $y_3 = k$. Moreover, the minimum kurtosis for any pair of constraints in the k-line will be $\sigma_{p^4} = \overline{\sigma}_{p^4} R^4$, or $y_{p^4} = \overline{y}_{p^4} \|R\|$.

The consequence of the above proposition is that exactly the same homothecy applies in 3D space defined by the standardised kurtosis axis and the two odd-moments axes. The results in Proposition 2 are then easily translated to the present context and the following is valid as well:

Proposition 3*: The direction in the R x y_3 half plane that gives the highest R for all the minimum kurtosis portfolios with the same standardised kurtosis y_4 is unique. Moreover, in this direction, the skewness constraint to programme (10) is not binding. As regards skewness, there is at least one direction giving the maximum skewness, where the constraint property applies.

The solution to the problem of minimising kurtosis for a given excess return is:

$$M_4(\alpha_R \otimes \alpha_R \otimes \alpha_R) = \frac{R}{B_{(-2)}} x \quad , \tag{14}$$

which, when $R = 1$, becomes: $$M_4(\overline{\alpha}_R \otimes \overline{\alpha}_R \otimes \overline{\alpha}_R) = \frac{1}{\overline{B}_{(-2)}} x \quad . \tag{15}$$

The systems of weights defined by $\alpha_R = \overline{\alpha}_R R$ are solutions to (12); thus one only needs to find one portfolio $\overline{\alpha}_R$ to generate the whole set of minimum kurtosis portfolios for a given R. The skewness corresponding to α_R is given by:

$$\sigma_{R^3} = \frac{B_0 R}{B_{(-2)}} = \frac{B_0}{\overline{B}_{(-2)}} R^3 \quad ; \tag{16}$$

so that the angular coefficient

$$k_R = \left(\frac{B_0}{\overline{B}_{(-2)}} \right)^{1/3} \tag{17}$$

defines a direction in the expected returns x skewness plane which is the "maximum mean returns line" for a given (minimum) kurtosis.

The "maximum mean returns line" divides the minimum iso-kurtosis curves in two parts; since agents want the highest possible skewness, they will probably work with the upper half of the curve. In contrast to the classical case of minimising variance for a given return, there is no closed form for the portfolio weights $\overline{\alpha}_R$, as can be seen from (15). However, it is possible to show that this function is strictly convex in its entire domain, therefore implying that the solution is unique.

The highest/lowest skewness directions, as in the case of variance, will be the ones associated to the solution of the problem of finding the lowest kurtosis subject to a given skewness. Calling these portfolios α_S, they are implicitly defined by system

$$M_4(\alpha_s \otimes \alpha_s \otimes \alpha_s) = \frac{\sigma_{s^3}}{B_2} M_3[\alpha_s \otimes \alpha_s] \quad , \tag{18}$$

the portfolio that solves the problem when $\sigma_{p^3} = 1$ is naturally defined by:

$$M_4(\bar{\alpha}_s \otimes \bar{\alpha}_s \otimes \bar{\alpha}_s) = \frac{1}{\bar{B}_2} M_3[\bar{\alpha}_s \otimes \bar{\alpha}_s] \quad . \tag{19}$$

It can also be easily verified that the mean return related to the solution of (18) is

$$R = \frac{\sigma_{s^3}}{B_2} B_0 \quad , \tag{20}$$

so that the directions are defined by

$$k_s = \left(\frac{B_0}{\bar{B}_2} \right)^{-1/3} \quad . \tag{21}$$

Unfortunately, in this case, there can be more than one solution, and consequently more than one direction with a local maximum skewness for a given level of kurtosis. Notwithstanding, the projection of each iso-kurtosis curve will also be enveloped, in the first quadrant, by the two axes and two tangent lines parallel to them.

3. Generalising for higher even moments.

We now consider the general case of minimising an even moment given the two first odd moments. The lagrangian of the problem will be:

$$\alpha' M_p \alpha^{\otimes(p-1)} + \lambda(R - \alpha' x) + \gamma(\sigma_{p^3} - \alpha' M_3 \alpha^{\otimes 2}) \quad , \tag{22}$$

giving the first order conditions:

$$\begin{aligned} &pM_p \alpha^{\otimes(p-1)} = \lambda x + 3\gamma M_3 \alpha^{\otimes 2} \\ &R = \alpha' x \\ &\sigma_{p^3} = \alpha' M_3 \alpha^{\otimes 2} \quad . \end{aligned} \tag{23}$$

Noticing that $M_p \alpha^{\otimes(p-1)} = M_p(\alpha^{\otimes(p-2)} \otimes I_n)\alpha$, and that matrix $M_p(\alpha^{\otimes(p-2)} \otimes I_n)$ is symmetric and positive definite, the following system can be formed from (23) to give the values of the multipliers:

$$
\begin{aligned}
pR &= \lambda x'(M_p \alpha^{\otimes(p-2)} \otimes I_n)^{-1} x + 3\gamma x'(M_p \alpha^{\otimes(p-2)} \otimes I_n) M_3 \alpha^{\otimes 2} \\
p\sigma_{p^3} &= \lambda (M_3 \alpha^{\otimes 2})'(M_p \alpha^{\otimes(p-2)} \otimes I_n)^{-1} x + 3\gamma (M_3 \alpha^{\otimes 2})'(M_p \alpha^{\otimes(p-2)} \otimes I_n) M_3 \alpha^{\otimes 2} \quad . \quad (24)
\end{aligned}
$$

Defining

$$B_{2-p} = x' \left[M_p(\alpha^{\otimes(p-2)} \otimes I_n) \right]^{-1} x \quad ,$$

$$B_{4-p} = x' \left[M_p(\alpha^{\otimes(p-2)} \otimes I_n) \right]^{-1} M_3 \alpha^{\otimes 2} \quad \text{and}$$

$$B_{6-p} = (\alpha^{\otimes 2})' M_3' \left[M_p(\alpha^{\otimes(p-2)} \otimes I_n) \right]^{-1} M_3 \alpha^{\otimes 2} \quad ,$$

with the subscripts corresponding to the generalised degree of homogeneity with respect to the vector of weights, the final solution comes from the system:

$$(B_{2-p} B_{6-p} - B_{4-p}{}^2) \ M_p \alpha^{\otimes(p-1)} = (B_{6-p} R - B_{4-p} \sigma_{p^3}) x + (B_{4-p} R - B_{2-p} \sigma_{p^3}) M_3 \alpha^{\otimes 2} \quad ; \quad (25)$$

the optimal portfolio p-th moment being:

$$\sigma_{p^p} = \frac{B_{6-p} R^2 - 2 B_{4-p} R \sigma_{p^3} + B_{2-p} (\sigma_{p^3})^2}{B_{2-p} B_{6-p} - B_{4-p}{}^2} \quad . \quad (26)$$

Again, the similarities (2)-(12)-(25) and (4)-(13)-(26) should be stressed.

The following result summarises all the properties of the solutions set:

Theorem: For a given p=2,4, ..., consider in (R, y_3, y_p) space of standardised moments a iso-p-th moment curve Γ of solutions to (22)

THEN

i) the optimal portfolios set is contained in the cone $\{O\} * \Gamma$, where O=(0,0,0) is the origin of (R, y_3, y_p) space;

ii) the projection of Γ in the Rxy_3 plane is a curve: a) symmetric to the origin and b) inscribed in a rectangle whose sides are parallel to the axes; the vertical and horizontal sides correspond, resp., to the highest (and lowest) R and y_3 values which produce a solution in Γ.

Proof (we outline the steps of the proof)**:** For proving i) one first follows steps similar to those in Propositions 1 and 1*, showing that on each line passing through the origin and a general point (R,y_3), the solutions to (22) increase linearly either with R – if the solution to $(1, y_3/R)$ is taken as the fundamental one – or with y_3 – if the solution to $(R/y_3, 1)$ is the one fixed. As the origin O=(0,0,0) solves (22), this is sufficient to demonstrate that any solution will be in the cone. In the case of ii), the symmetry is seen by the fact that reverting to the pair $(-R,-y_3)$ does not change either (25) or (26). As regards the tangents, a reasoning similar to the ones in the previous section determines the points relative to the highest R and y_3, by symmetry the points of the lowest R and y_3 are obtained and the rectangle can be traced. □

This basic result is important in finding the efficient portfolios set for the three moments at stake. It is easy to convince oneself that not all points in the cone will characterise an efficient portfolio, though, of course, the efficient set will be contained in the cone (see Athayde and Flôres (2001)). Moreover, one could be tempted to derive the following

(false) **Corollary:** If problem (22) has a solution THEN the optimal value is unique.

Indeed, by the Theorem, if (22) has a solution then the optimal p-th moments must lie in the cone. They will be found in the intersection of a vertical line through the point defined by the given odd moments in the Rxy_3 plane and the cone. Simple properties of a cone in finite dimensional Euclidian spaces ensure that this intersection is unique.

This nice property would mean that knowledge of the geometric structure of the optimal portfolios set had allowed a simple and elegant proof of uniqueness. However, such an argument would be circular, as the curve Γ used to characterise the cone is supposedly the curve formed already by the minimum p-th moments, related to the optimal solutions of (25). It is worth reminding that system (25), as its special cases (2) and (12), *implicitly* defines the optimal weights, and may as well have more than one solution. These others either will be local, not global optima or it might even happen that different optimal vectors α yield the same optimal p-th moment in (26). Propositions 1 to 3 (and 1* and 3*) are valid for *any of these solutions* – this meaning that even different "solution cones" may exist; but the Theorem considers, by hypothesis, *the* "optimal cone", and so the Corollary is senseless. Unfortunately, at the present stage, we do not have a general, deeper knowledge of the structure of the solutions set. Moreover, the hypothesis also requires the existence of a solution; rigorous conditions for guaranteeing this, as regards system (25), are still an open question.

An interesting special case of (22) is when only a mean return restriction is imposed, the skewness constraint being disregarded. Without much difficulty one sees that the first order conditions become:

$$pM_p\alpha^{\otimes(p-1)} = \lambda x$$

$$R = \alpha' x \qquad . \qquad (27)$$

So that the optimal weights must solve the system

$$(B_{2-p}) \quad M_p\alpha^{\otimes(p-1)} = Rx \qquad ; \qquad (28)$$

and the corresponding p-th moment bears the following relationship with the given return:

$$\frac{\sigma_{p^p}}{R^2} = (B_{2-p})^{-1} \qquad . \qquad (29)$$

In this case, the homothecy property implies that *only one system needs to be solved*, namely, the one obtained by setting $R=1$ in (28).

4. Further properties and extensions.

In order to give a further insight both on the geometric aspects discussed as well as on the difficulties involved in the solution of system (28), we consider the special problem of minimising kurtosis given expected return, in the case of two assets and setting to zero all co-kurtoses where an asset appears only once. This leaves us with three distinct non-zero elements in the kurtosis tensor, and the M_4 matrix – shown, in the general case, in the Appendix – becomes:

$$\begin{bmatrix} \sigma_1 & 0 & 0 & \sigma_{12} & 0 & \sigma_{12} & \sigma_{12} & 0 \\ 0 & \sigma_{12} & \sigma_{12} & 0 & \sigma_{12} & 0 & 0 & \sigma_2 \end{bmatrix} \quad .$$

The simplified notation used for the subscripts, suppressing repetition of identical indexes, stresses the identical values and should cause no confusion. Notice that, unless the assets distributions are singular, all entries are strictly positive.

Calling $\alpha = (\alpha_1, \alpha_2)'$ the vector of weights, and noticing that:

i) $M_4 \alpha^{\otimes 3} = \begin{bmatrix} \alpha_1^3 \sigma_1 + 3\alpha_1 \alpha_2^2 \sigma_{12} \\ 3\alpha_1^2 \alpha_2 \sigma_{12} + \alpha_2^3 \sigma_2 \end{bmatrix}$;

ii) matrix $\left[M_4 (\alpha^{\otimes 2} \otimes I_2) \right]^{-1}$ will be equal to:

$$\Delta^{-1} \begin{bmatrix} \alpha_1^2 \sigma_{12} + \alpha_2^2 \sigma_2 & -2\alpha_1 \alpha_2 \sigma_{12} \\ -2\alpha_1 \alpha_2 \sigma_{12} & \alpha_1^2 \sigma_1 + \alpha_2^2 \sigma_{12} \end{bmatrix}$$

where $\Delta = \alpha_1^4 \sigma_1 \sigma_{12} + \alpha_1^2 \alpha_2^2 (\sigma_1 \sigma_2 - 3\sigma_{12}^2) + \alpha_2^4 \sigma_{12} \sigma_2$ is the determinant of the direct matrix;

one is ready to build up system (28). Of course, as said in the previous section, only one solution matters, namely that which considers $R=1$. We shall, however, impose the additional assumptions that the marginal kurtoses are equal (i.e., $\sigma_1 = \sigma_2 = \sigma$) and that excess returns for both assets are also equal (to a common value x). With this, we can finally write system (30):

$$\begin{aligned} &[\alpha_1^5 (\sigma \sigma_{12} + 2\sigma^2) - 4\alpha_1^4 \alpha_2 \sigma \sigma_{12} + \alpha_1^3 \alpha_2^2 (3\sigma_{12}^2 + 7\sigma \sigma_{12}) - 12\alpha_1^2 \alpha_2^3 \sigma_{12}^2 + 3\alpha_1 \alpha_2^4 \sigma_{12}^2] x = \\ &\qquad = \alpha_1^4 \sigma\, \sigma_{12} + \alpha_1^2 \alpha_2^2 (\sigma^2 - 3\sigma_{12}^2) + \alpha_2^4 \sigma\, \sigma_{12} \end{aligned}$$

$$\begin{aligned} &[3\alpha_1^4 \alpha_2 (2\sigma \sigma_{12} + \sigma_{12}^2) - 12\alpha_1^3 \alpha_2^2 \sigma_{12}^2 + \alpha_1^2 \alpha_2^3 (3\sigma_{12}^2 + \sigma \sigma_{12} + 2\sigma^2) - 4\alpha_1 \alpha_2^4 \sigma \sigma_{12} + \alpha_2^5 \sigma \sigma_{12}] x = \\ &\qquad = \alpha_1^4 \sigma\, \sigma_{12} + \alpha_1^2 \alpha_2^2 (\sigma^2 - 3\sigma_{12}^2) + \alpha_2^4 \sigma\, \sigma_{12} \end{aligned}$$

Given the symmetry of the parameter values, the optimal weights will be identical, being easy to see that their common value is:

$$\alpha = \frac{1}{2x} \qquad . \tag{31}$$

These weights, however, can be related to either maxima or minima. For the latter, the bordered Hessian sufficient condition[5] amounts, in this case, to check whether matrix

$$\begin{bmatrix} 12\alpha^2(\sigma+\sigma_{12}) & 24\alpha^2\sigma_{12} & -x \\ 24\alpha^2\sigma_{12} & 12\alpha^2(\sigma+\sigma_{12}) & -x \\ -x & -x & 0 \end{bmatrix} \tag{32}$$

has a negative determinant. Replacing α by its value in (31), the condition becomes:

$$6(\sigma_{12}-\sigma)<0 \qquad or \qquad \sigma_{12}<\sigma \qquad . \tag{33}$$

The symmetric weights solution produces a minimum *only if* the non-null co-kurtosis is smaller than the common marginal kurtosis.

This rather simple example may serve as an illustration of how far intuition can help when considering higher moments, as well as of the impact of simplifications in the higher-moments tensors. The final solution is independent of the marginal kurtoses *and* of the even co-kurtosis. Indeed, as the risk measures have a completely symmetric structure as regards the two (risky) assets, the identical weights can be found by direct solution of the excess return constraint. The higher the identical return, obviously the less will be purchased of each risky asset – as the portfolio excess return is *fixed in 1* – and more will be put in the riskless asset[6].

Given the similar roles played by kurtosis and variance, we could then expect that the same would apply for the identical weights that result when equal marginal variances are used instead of kurtosis. In fact, (31) is exactly the solution to (5) in this case, the (common) variances and co-variances playing no role at all. Moreover, use of the bordered Hessian condition shows that a minimum exists only if

$$\sigma_{12}<\sigma \qquad . \tag{34}$$

Though "identical" to (33), (34) will be always valid if the assets covariance is negative, what cannot happen in the case of the even co-kurtosis. Indeed, in our simplified kurtosis context, there is no room for diversification.

Absent from (31) – in its two versions/solutions -, the risk measures do however play a role. Beyond determining whether a minimum has been achieved, they explicitly

---

[5] See, for instance, Theorem 9.9, page 202, in Panik (1976).
[6] Asymptotically, all the weight will go to the riskless asset.

appear in the shadow price of the restrictions, given by the value of the Lagrange multipliers. These are equal to $\lambda = \frac{\sigma + \sigma_{12}}{2x^2}$ in the variance case, and to $\lambda = \frac{\sigma + \sigma_{12}}{2x^2}$ in that of kurtosis[7]. The formal identity of the two values hides different behaviours. Again, in the case of the second moment, a negative covariance may substantially decrease "the cost" of the unit return restriction. On the other hand, both (nonnegative) kurtoses add up, penalising more heavily an increase in the fixed return.

Summing up, the example shows that the choice to minimise either kurtosis or variance (in this very simple, symmetric case) has, in spite of producing exactly the same solution weights, fairly different implications. Moreover, radical simplifications in the moments tensor may produce rather particular solutions. A small change in the example, like allowing for different marginal kurtoses, would completely alter the above discussion. Informally speaking, introducing higher moments in portfolio choice makes it a "more non-linear" problem and, consequently, much more sensible to small changes in the initial conditions.

5. Concluding remarks.

The availability of a general method to treat portfolio choice in a higher moments context seems an unquestionable advantage. We outlined in the previous sections one such method, that allows for a compact, analytical treatment of all formulas involved in the optimisation problem. Thanks to this, powerful geometric insights could be gained.

Nevertheless, the task before anyone interested in the subject is still nearly formidable. A basic existence result and more insights on the solutions set would be welcome. Final characterisation of the efficient portfolios set requires more than the techniques here discussed, duality methods being needed to completely identify the efficient points. We solved this up to the fourth moment, Athayde and Flôres (2001, 2002), but a general method seems possible. Moving from static to dynamic optimisation frameworks generates additional, rather difficult theoretical and computational problems[8].

---

[7] The reader should keep in mind that both σ and σ_{12} have different meanings in the two formulas.
[8] Work in this direction has been initiated with Berç Rustem (Imperial College, London).

Last, but not least, as section 4 glimpsed into, the number of different situations in the higher moments case is extremely large, a great probability existing of senseless or unattractive special formulations. These can only be sorted out through a combination of more theoretical findings with several examples and applied experiments. The notation developed, and its corresponding algebra, may help in designing many of these experiments.

Appendix: The matrix notation for the higher moments arrays.

Dealing with higher moments can easily become algebraically cumbersome. Given a n-dimensional random vector, the set of its p-th order moments is, as a mathematical object, a tensor. The second moments tensor is the popular *nxn* covariance matrix, while the third moments one can be visualised as a *nxnxn* cube in three-dimensional space. However, the tensor notation, which is so useful in physics, geometry and some areas of statistics (see, for instance, McCullagh (1987)), did not appear so convenient to deal with the portfolio choice problem. We then developed a special notation, which allows performing all the needed operations within the realm of matrix calculus. The advantages of this are manifold. Beyond having a synthetic way to treat complicated expressions, the mathematical tools required are standard linear algebra results and, with the help of Euler's theorem – as most real functions involved are homogeneous in the vector of portfolio weights -, a differential calculus easily ensues. Moreover, the different formulae and systems arrived at are written in a compact and straightforward way, easily translated into formal programming languages.

Before presenting the notation, we remind that, throughout the paper we deal with *all* the possible p-moments of a given n-dimensional random vector of asset returns. Undoubtedly, the difficulty in manipulating all these values simultaneously has been a deterrent to tackle the problem in its full generality. Thinking of skewness and kurtosis, for instance, the respective three and four-dimensional "cubes", where several identical values are found, have n^3 and n^4 elements. Of course, in practice, gathering all these values may quickly become a formidable task. Indeed, as an example, the number of *different* kurtoses is in principle $\binom{n+3}{4}$, what, in the case of five assets, gives already 70 values to be computed. It is then very likely that, in each practical problem, either a significant number of co-moments will be set *a priori* to zero or another simplifying assumption will be used, and very seldom one will work with the full set of cross moments. However, as said in the introduction, the great variety of possible assumptions is an extra argument for a general treatment of the problem.

Our notation transforms the full p-th moments tensor, with n^p elements, into a matrix of order nxn^{p-1} obtained by slicing all bi-dimensional nxn^{p-2} layers defined by fixing one asset and then taking all the moments in which it figures at least once and pasting them, in the same order, sideways. Row *i'* of the matrix layer corresponding to have fixed the *i-th* asset gives – in a pre-established order – all the moments in which assets *i* and *i'* appear at least once. Of course, assets must be ordered once and for all and this order respected in the sequencing of the layers and in the numbering of the rows of each layer. Accordingly, a conformal ordering must be chosen, and thoroughly used, for the combinations (with repetitions) of n elements into groups of p-2 that define the columns of each matrix layer.

In the case of kurtosis, for instance, two indices/variables/co-ordinates must be held constant. Calling σ_{ijkl} a general (co-) kurtosis, when n=2, the resulting 2x8 matrix will be:

$$\begin{bmatrix} \sigma_{1111} & \sigma_{1112} & \sigma_{1121} & \sigma_{1122} & \sigma_{1211} & \sigma_{1212} & \sigma_{1221} & \sigma_{1222} \\ \sigma_{2111} & \sigma_{2112} & \sigma_{2121} & \sigma_{2122} & \sigma_{2211} & \sigma_{2212} & \sigma_{2221} & \sigma_{2222} \end{bmatrix}$$

where, as expected, many entries are identical.

Now suppose that a vector of weights $\alpha \in R^n$ is given, and M_1, M_2, M_3, ... and M_p stand for the matrices containing the expected (excess) returns, (co-)variances, skewnesses ... and p-moments of a random vector of *n* assets. The mean return, variance, skewness ... and p-th moment of the portfolio with these weights will be, respectively: $\alpha' M_1$, $\alpha' M_2 \alpha$, $\alpha' M_3 (\alpha \otimes \alpha)$ ... and $\alpha' M_p (\alpha \otimes \alpha \otimes \alpha \ldots \otimes \alpha) \equiv \alpha' M_p \alpha^{\otimes p-1}$ where '⊗' stands for the Kronecker product.

The above expressions provide a clue on the mentioned advantages of the notation. The fact that the tensors were transformed into matrices allows the use of matrix algebra – and differential calculus - in all expressions and derivations, giving way to compact and elegant formulas. It is immediate to see that, as real functions of α, the four expressions above are homogenous functions of the same degree as the order of the corresponding moment. This means that Euler's theorem can be easily used in the needed derivations.

As an example, the derivative of the portfolio kurtosis with respect to the weights will be:

$$\frac{\partial}{\partial\alpha}[\alpha' M_4(\alpha\otimes\alpha\otimes\alpha)] = 4M_4(\alpha\otimes\alpha\otimes\alpha) = 4M_4\alpha^{\otimes 3} \qquad .$$

References

Adcock, C. J. 2002. Asset pricing and portfolio selection based on the multivariate skew-student. *This volume.*

Adcock, C. J. and K. Shutes. 1999. Portfolio selection based on the multivariate skew-normal distribution, in A. Skulimowski, ed., *Financial Modelling,* Progress and Business Publishers, Krakow.

Athayde, G. M. and R. G. Flôres Jr. 1997. *A CAPM with Higher Moments: Theory and Econometrics.* EPGE/FGV, Ensaios Econômicos n° 317, Rio de Janeiro.

Athayde, G. M. and R. G. Flôres Jr. 2000. Introducing higher moments in the CAPM: some basic ideas, in C. L. Dunis, ed., *Advances in Quantitative Asset Management,* Kluwer Academic Publishers, Norwell, Mass.

Athayde, G. M. and R. G. Flôres Jr. 2001. Finding a maximum skewness portfolio – a general solution to three-moments portfolio choice. *Journal of Economic Dynamics and Control, forthcoming.*

Athayde, G. M. and R. G. Flôres Jr. 2002. A journey through an undiscovered country: efficient portfolio sets with four moments. *Submitted.*

Barone-Adesi, G. 1985. Arbitrage equilibrium with skewed asset returns. *Journal of Financial and Quantitative Analysis* 20, 299-313.

Harvey, C. R. and A. Siddique. 1999. Auto-regressive conditional skewness. *Journal of Financial and Quantitative Analysis* 34 (4), 465-87.

Harvey, C. R. and A. Siddique. 2000. Conditional skewness in asset pricing tests. *The Journal of Finance* 60, 1263-95.

Hwang, S. and S. Satchell. 1999. Modelling emerging market risk premia using higher moments. *International Journal of Finance and Economics* 4 (4), 271-96.

Jurczenko, E. and B. Maillet. 2001. The 3-CAPM: Theoretical foundations and an asset-pricing model comparison in a unified framework, in C. L. Dunis, A. Timmerman and J. Moody, eds., *Developments in Forecast Combination and Portfolio Choice*, John Wiley & Sons.

Jurczenko, E. and B. Maillet. 2002. The four-moment capital asset pricing model: some basic results. *This volume.*

McCullagh, P. 1987. *Tensor Methods in Statistics.* Chapman and Hall, London.

Panik, M. J. 1976. *Classical Optimization: Foundations and Extensions.* North-Holland Publishing Company, Amsterdam.

Pedersen, C. S. and S. E. Satchell. 1998. An extended family of financial risk measures. *Geneva Papers on Risk and Insurance Theory* 23, 89-117.

Pedersen, C. S. and S. E. Satchell. 2000. Small sample analysis of performance measures in the asymmetric response model. *Journal of Financial and Quantitative Analysis* 35, 425-50.

Scott, R. and P. A. Horvath. 1980. On the direction of preference for moments of higher order than the variance. *The Journal of Finance* 35, 915-19.